Branching processes

Let $Z_{i1}, Z_{i2}, \ldots, i=1, 2, \ldots$ be iid counting r.v.'s with pgf $G$. Define the stochastic process
\[
\{X_0, X_1, \ldots\} \text{ by } X_0 = 1 \text{ and } X_m = Z_{m-1,1} + \cdots + Z_{m-1, X_{m-1}}.
\]
The process $\{X_t \mid t = 0, 1, \ldots\}$ is called a Galton-Watson branching process with offspring pgf $G$.
Denote the offspring mean $\mu$ and variance by $\mu$ and $\sigma^2$ respectively
\[
(\mu = E(Z_{i1}), \sigma^2 = Var(Z_{i1})).
\]
Branching processes are classified as subcritical, critical or supercritical as $\mu < 1$, $\mu = 1$ or $\mu > 1$. 

Remarks

- Clearly $X_t | X_{t-1} \equiv X_t / X_{t-1}$, so that the process is Markov. Also,$X_{t+1} | X_t \equiv X_1 / X_0$, so that the process is time homogeneous.

- The assumption $X_0 = 1$ is for convenience and is easily dropped.

- $X_t$ is the population at the $t$-th generation. Individuals live for one generation and have offspring according to the offspring dist'n.

Set $G(t) = E(\ell X_t)$, $\rho_t = P(X_t = 0)$, $p = \lim_{t \to \infty} \rho_t$ and $G(t)(s) = t$-th iterate of $G$. 
We then note that $p_\infty$ increases and since it is bounded the limit $p_\infty$ exists. It is called the probability of ultimate extinction. We also have

$$G_m(s) = G[G_{m-1}(s)]$$

which incidentally yields

$$G_m(s) = G_m(s)$$

So

$$p_m = P(X_m = 0) = G_m(0) = G[G_{m-1}(0)] = G(p_{m-1})$$

and hence

$$p = G(p)$$

Now notice $G'(s)$ and $G''(s) > 0$ on $s > 0$ so that $G$ is strictly increasing and convex on $s > 0$. 
If $\mu = G'(1) \leq 1$ then there will only be one root of 
\[ p = G(p) \quad ; \quad 0 \leq p \leq 1 \]
and that is at $p=1$ (certain extinction).

If $\mu > 1$ then there will be 2 roots - one at $p=1$ and the other at $p<1$. It is clear that $p$ must be the smaller of the 2 roots. Indeed, call the smaller root $p_0$. Then we have
\[ p_1 \leq p_0 \Rightarrow p_2 = G(p_1) \leq G(p_0) = p_0 \Rightarrow \ldots \Rightarrow p_m \leq p_0 \]
so that $\lim p_m = p_0$.

Notice that $E(X_t) = \mu^t$. If
the process does not die off \( \mu > 1 \) case and condition on non-extinction) then it can be shown \( X_m \to 0 \). Set

\[
W_m = \frac{X_m}{E(X_m)}
\]

so that \( E(W_m) = 1 \). Use past to show

\[
Var(W_m) = \frac{\sigma^2 (1 - \mu^{-m})}{\mu^2 - \mu}
\]

so that

\[
Var(W_m) \to \frac{\sigma^2}{\mu^2 - \mu} < 0
\]

In fact we can show \( W_m \to W \) using martingale arguments. We return to this matter latter.
A bit on recurrence/renewal

Let $X_1, X_2, \ldots$ be independent rv's $\in \mathbb{N} = \{1, 2, \ldots\}$. Furthermore, take $X_2, X_3, \ldots$ to be i.i.d with p.d.f $G$ and set $D(x) = E(e^{xt})$. All times $T_k = X_1 + \cdots + X_k$, $k = 1, 2, \ldots$ an "event" $H$ occurs. At other times it does not. The $X_i$'s are called interarrival/interoccurrence times.

Let $H_n = \mathbb{1}_H$ occurs at time $n \frac{3}{2}$ & $\nu_n = P(H_n)$. 

\[
\begin{array}{c}
X_1 \xrightarrow{X} X_2 \xrightarrow{X} X_3 \xrightarrow{X} \ldots \\
0 \xrightarrow{H} X \xrightarrow{H} X \\
\end{array}
\]
Let
\[ U(a) = \sum_{m=1}^{\infty} u_m a^m \]
be the generating function of \(u_0, u_1, \ldots\).

This is not a pdf! We can show
\[ U(a) = \frac{D(a)}{1-G(a)} \]
which allows us to calculate \(u_m\), at least in principle. Consider, for example, set
\[ D_\star(a) = \frac{1-G(a)}{\mu (1-a)} \quad \text{where} \quad |a| < 1 \]
where \(\mu = E(X_2)\) is the mean interarrival time. Let
\[ D_\star(1) = \lim_{\Delta t \to 1} D_\star(a) = -\frac{G'(1)}{-\mu} = 1 \]
Now expand $D_*(s)$ in a power series and notice that the coefficient of $s^m$ is
\[d_m = \mu^{-1} \left[ 1 - (P(X_2=1) + \ldots + P(X_2=m)) \right].\]
\[= P(X_2 > m) / \mu \quad (\mu > 0).\]

Since
\[E(X_2) = \sum_{m=0}^{\infty} P(X_2 > m),\]
we conclude
\[D_*(s) = \sum_{m=0}^{\infty} d_m s^m\]
is a pgf (since $d_m > 0$ and $\sum d_m = 1$).
So in this case
\[U(s) = \frac{1}{\mu (1-s)}\]
which yields $u_m = \frac{1}{\mu} \quad ; \quad m=1, 2, \ldots$
which is a constant sequence.
If we take $X_i \sim D_\star$ notice that $X_i \in \mathbb{N} \cup \{0\} = \mathbb{Z}^+$ then

$$u_m = \frac{1}{m} \to \frac{1}{\mu}$$

This suggests $u_m \to V_\mu$ more generally since the initial distribution's influence will fade in time. This result is the renewal theorem.

**Renewal Theorem**

If $\mu$ is finite and $\gcd\{m : P(X_\star = m) > 0\} = 1$

then

$$u_m \to V_\mu$$

This is an important result and may be proved in a variety of ways. A modern approach is to use "coupling". For now we show

$$U(s) = \frac{D(s)}{\lambda(s)}$$

$$U(s) = D(s) \cdot [1 - G(s)]$$
Suppose \( X_1, X_2, \ldots \) are i.i.d. \( \epsilon \in \{1, 2, \ldots\} \)

Let the common pdf be

\[
G(a) = E(A_{X_1}) = E(A_{X_2}) = \ldots
\]

Now let \( u_m = P(\text{renewal at time } m) \) \( \epsilon \)

\[
U(a) = \sum_{m=1}^{\infty} u_m a^m,
\]

which is the generating function.

Now let \( T_m \) be the time of the \( m \)th renewal.

So \( T_m = X_1 + \cdots + X_m \)

Look at \( H_m = \{\text{renewal at time } m\} \)

its indicator is \( I(\mathbb{1}(H_m)) \).

\[
I(\mathbb{1}(H_m)) = 1, \text{ if a renewal at time } m
\]

\[
= 0, \text{ otherwise}
\]

Clearly \( \mathbb{E}[I(\mathbb{1}(H_m))] = u_m \)
Now
\[
\sum_{m=1}^{\infty} I(H_m) s^m
\]
converges at least on \(|s| < 1\),
\[
E\left[\sum_{m=1}^{\infty} I(H_m) s^m\right] = \sum_{m=1}^{\infty} E[I(H_m)] s^m
\]
\[
= \sum_{m=1}^{\infty} w_m s^m = L(s)
\]

Now look at
\[
\sum_{m=1}^{\infty} a T_m
\]
which also converges for \(|a| < 1\),
\[
E\left(\sum_{m=1}^{\infty} a T_m\right) = \sum_{m=1}^{\infty} E(a T_m)
\]
\[
= \sum_{m=1}^{\infty} E(a x_1 + x_2 + \cdots + x_m)
\]
\[
= \sum_{m=1}^{\infty} E(a x_1) E(a x_2) \cdots E(a x_m)
\]
\[
\sum_{m=1}^{\infty} G(x)^m = \frac{G(x)}{1 - G(x)} \quad (D(x) = \frac{D(x)}{1 - G(x)}, \quad \text{if} \quad D(x) = E(g(x)))
\]

But
\[
\sum_{m=1}^{\infty} I(H_m) \Delta^m = \sum_{m=1}^{\infty} A \Delta^m
\]

\[\hat{\text{This is always true even if } X_1 \text{ has a different distribution than } X_2, X_3, \ldots.}\]

\[\therefore U(x) = \frac{G(x)}{1 - G(x)}\]
The Renewal Theorem

Let \( X_i \in \{0, 1, 2, \ldots \} \) be independent of \( X_2, X_3, \ldots \), which are i.i.d. nonperiodic counting r.v.'s \( > 0 \) with pgf \( G \). Then

\[ u_m \to \frac{1}{\mu} \]

Proof: Define a \(*\)-renewal process such that \( X_1^*, X_2^*, \ldots \) are iid of the \( X_i \)'s; \( X_1^* \) has pgf \( D_1 + X_2^*, X_3^*, \ldots \) are iid with pgf \( G \). We know that

\[ u_m^* \to \frac{1}{\mu} \]

Now let \( T \) be the first time where \( H \) occurs (simultaneously) in both processes. "Clearly" \( T < \infty \).

After \( T \) the \(*\) process + the original process are indistinguishable. So
\[ u_m = P(H_m \mid T \leq m) P(T \leq m) \\
+ P(H_m \mid T > m) P(T > m) \]

\[ = P(H_m^* \mid T \leq m) P(T \leq m) \\
+ P(H_m \mid T > m) P(T > m) \]

\[ \hat{H}_m \text{ is the event that } H \text{ occurs at time } m \text{ for the original process } \]

\[ \text{refers to the } * \text{ process!} \]

Also,

\[ \frac{1}{\mu} = u_m^* = P(H_m^* \mid T \leq m) P(T \leq m) \\
+ P(H_m^* \mid T > m) P(T > m) \]

\[ |u_m - u_m^*| = P(T > m) |P(H_m \mid T > m) - P(H_m^* \mid T > m)| \]

\[ \lesssim 2 P(T > m) \rightarrow 0 \text{ as } m \rightarrow \infty. \]

\[ u_m \rightarrow \frac{1}{\mu} \]

\[ \text{qed} \]
Renewal processes

Let \( X_1, X_2, \ldots \) be iid \( \geq 0 \) rvs.

Set

\[
S_m = \sum_{i=1}^{m} X_i, \quad m \geq 1
\]

and \( S_0 = 0 \). The \( S_m, m \geq 1 \), will be the times of renewals (or events) and if \( N(t) = \# \) of renewals in \((0, t]\) then \( \{N(t), t \geq 0\} \) will be termed a renewal (counting) process. By convention we will not start with a renewal \( (N(0)=0) \). Also it will be convenient to allow \( X_1 \) to have a different distribution than \( X_i, i \geq 2 \). In this case we have a delayed renewal process. \( \mu = E(X_m) \) is the mean interarrival time (the \( X \)'s are called interarrival times). We will assume \( P(X_m=0) < 1 \) so that \( \mu > 0 \).
The SLLN gives
\[ \frac{S_m}{m} \xrightarrow{a.s.} \mu \]
so that \( S_m \xrightarrow{a.s.} \infty \). This tells us that \( N(t) \) is finite (w.p.1), but we will drop this qualification when there can be no confusion - the same approach was taken with \( E(Y|X) \).

\[ m(t) = E[N(t)] \]
is called the renewal function.

Incidentally, we have already seen a renewal process, namely the Poisson process. There the \( X \)'s were i.i.d. exponentials and \( m(t) = \lambda t \), where \( \lambda \) was the rate.
Clearly

\[ N(t) \geq m \iff S_m \leq t \]

so that

\[ P(N(t) \geq m) = P(S_m \leq t) \]

\( S_m \) is a sum of iid rv's so that its df, \( F_m \) say, is the \( m \)-fold convolution of \( F \) (the df of the X's) with itself.

Remark. \( \text{If } X \text{ has df } F \) and \( Y \) has df \( G \), \( X, Y \) are independent then the distribution of \( X + Y \) is a convolution. In terms of df's it is denoted by \( F \ast G \) so that

\[ F \ast G (a) = P(X + Y \leq a) \]

(\( = \mathbb{E}_G (F(a - Y)) \) or \( \int F(a - y) dG \) means same)
Notice
\[ P[N(t) = m] = P(S_{m} \leq t) - P(S_{m+1} \leq t) \]

Denote the df of \( S_{m} \) by \( F_{m} \)
\( (= F \ast F \ast \ldots \ast F) \).

**Proposition**
\[ m(t) = \sum_{m=1}^{\infty} F(t) = \sum_{m=1}^{\infty} P(S_{m} \leq t) \]

**Proof**
\[ N(t) = \sum_{m=1}^{\infty} \mathbb{I}_{\{0 \leq S_{m} \leq t\}} \]
\[ \Rightarrow E(N(t)) = \sum_{m=1}^{\infty} E[\mathbb{I}_{\{0 \leq S_{m} \leq t\}}] \]
\[ = \sum_{m=1}^{\infty} P(S_{m} \leq t) \]
\[ = \text{ged} \]
**Proposition** \( m(t) < \infty \), \( t \geq 0 \)

**Proof** Since \( P(X_m > 0) \equiv \alpha \) it follows that \( P(X_m > \alpha) > 0 \). Now consider the new renewal process \( \tilde{N}_\alpha(t) \) with interarrivals

\[
X_{\alpha,m} = \begin{cases} 
0 & , \text{if } X_m < \alpha \\
\alpha & , \text{if } X_m \geq \alpha 
\end{cases}
\]

This new renewal process can have renewals only at times \( n\alpha \).

The \# of renewals at each of these times can be \( > 1 \) (since the \( X_m \)'s can be \( 0 \)). In fact, these numbers are independent geometric (\( P(X_m > \alpha) \)). Hence

\[
E[\bar{N}_\alpha(k\alpha)] = k / P(X_m > \alpha)
\]

(again \( N_\alpha(0) \equiv 0 \))
so that for $t > 0$

$$E[N_a(t)] \leq \frac{t/a}{P(X_m \geq a)}$$

Now since $X_{\alpha,m} \leq X_m$ we have $N_a(t) \geq N(t)$ so that $E(N(t)) < \infty$.

$qed$

Remarks.

(a) The same argument in fact shows $E(N_r(t)) < \infty$, for any $r > 0$!

(b) The $X_{\alpha,m}$ are rv's that can take on values on an equally spaced grid.

These are lattice rv's with the span being the largest spacing possible.
The span of a lattice rv is also called the period. For counting rv's (values in \( \mathbb{E}_0, 1, 2, \ldots 3 \)) we often use the terms aperiodic or non-arithmetic if the span is 1. This can lead to some confusion since lattice rv's are also called arithmetic rv's! For integer valued rv's we will try to use the term aperiodic to mean the span is 1 and periodic for spans >1. Unfortunately the terms non-arithmetic and arithmetic may creep in.
Set \( N(\infty) = \lim_{t \to \infty} N(t) \). Then

\[
P(N(\infty) < \infty) = P\left(X_m = \infty \text{ for some } m\right)
\]

\[
= P\left(\bigcap_{m=1}^{\infty} \{X_m = \infty\}\right)
\]

\[
\leq \sum_{m=1}^{\infty} P(X_m = \infty) = 0
\]

and so \( N(\infty) = \infty \).

\[\text{Proposition} \quad \frac{N(t)}{t} \xrightarrow{a.s.} \frac{1}{\mu}\]

\[\text{Proof:} \quad S_{N(t)} \leq t < S_{N(t) + 1}\]

\[
\begin{array}{c}
\text{time of renewal before or at } t \\
\downarrow \text{time of renewal after time } t
\end{array}
\]
Hence

\[ \frac{S_{N(t)}}{N(t)} \leq \frac{t}{N(t)} \leq \frac{S_{N(t)+1}}{N(t)} \]

Since \[ \frac{1}{N(t)} \to 0 \] we conclude

\[ \frac{N(t)+1}{N(t)} \xrightarrow{a} 1 \]

And

\[ \frac{S_{N(t)+1}}{N(t)} = \left( \frac{S_{N(t)+1}}{N(t)+1} \right) \frac{N(t)+1}{N(t)} \xrightarrow{a} \mu \]

As \[ \frac{S_{N(t)}}{N(t)} \xrightarrow{a} \mu \] we obtain

\[ \frac{t}{N(t)} \xrightarrow{a} \mu \]

Or

\[ \frac{N(t)}{t} \xrightarrow{a} \frac{1}{\mu} \]
So we have \( \frac{N(t)}{t} \to \frac{1}{\mu} \)

But this does not imply

\[ \frac{E(N(t))}{t} \to \frac{1}{\mu} \quad (\star) \]

However, in this case \((\star)\) is true which is a result known as the Elementary Renewal Theorem.

There are other versions such as

**Blackwell's Theorem**

(i) If the \( X \)'s are lattice rv's with period (i.e. span) \( d \), then

\[ E(\# \text{ of renewals at } md) \to d/\mu \]

(ii) If the \( X \)'s are not lattice rv's then

for \( a \geq 0 \)

\[ \frac{m(t+a) - m(t)}{a} \to \frac{a}{\mu} \]

Note: (i) holds for \( d = 1 \)
Notice that if we do not allow multiple renewals ($X_m > 0$) then
\[
E(\# \text{ of renewals at } m \text{d}) = P(\text{renewal at } m \text{d})
\]

In the case $d = 1$ this is just our $U_m$ which we proved $\rightarrow \frac{1}{\mu}$ using a coupling argument (there is still the issue of proving $T < \infty$).

If the $X$'s are non-lattice RV's and we assume
\[
\lim_{t \to \infty} m(t+a) - m(t) = g(a)
\]

exists then clearly
\[
g(a+b) = g(a) + g(b)
\]
so that $g(a) = c \cdot a$. Now use the Elementary Renewal Theorem to conclude $c = \frac{1}{\mu}$.
We will not prove Blackwell's Theorem in the non-lattice case. A CLT for renewal processes goes along the lines

\[ \frac{N(t) - t/\mu}{\sigma \sqrt{t/\mu^3}} \rightarrow N(0,1) \]

Here \( \sigma^2 = \text{Var}(X_1) \).

We now proceed to prove the Elementary Renewal Theorem.

Define \( X_1, X_2, \ldots \) be a sequence of rv's. A counting rv \( N > 0 \) is said to be a stopping time for the \( X_i \)'s if \( \{N \leq n\} \) is a function of \( X_1, \ldots, X_n \) for each \( n = 1, 2, \ldots \).
Let $X_1, X_2, \ldots$ be i.i.d. Bernoulli $(\frac{1}{2})$.

Set $N = \min \{ m : X_1 + \cdots + X_m = 10^3 \}$.

Then $N$ is a stopping time.

In fact $N$ is our familiar negative binomial.

Wald's Equation

Let $X_1, X_2, \ldots$ be i.i.d. with mean $\mu$ and let $N$ be a stopping time for the $X_i$'s such that $E(N) < \infty$. Then

$$E\left( \sum_{m=1}^{N} X_m \right) = E(N) \mu$$

Proof

$$\sum_{m=1}^{N} X_m = \sum_{m=1}^{\infty} X_m I_{\xi \geq m^3}$$

$$\Rightarrow E\left( \sum_{m=1}^{N} X_m \right) = \sum_{m=1}^{\infty} E(X_m I_{\xi \geq m^3})$$
\[= \sum_{m=1}^{\infty} E(X_m) \mathbb{E}\left(I_{N \geq m+3}\right), \quad \{N \geq m+3\} = \{N \geq m-1\}^c \]

\[= \mu \sum_{m=1}^{\infty} E\left(I_{N \geq m+3}\right) = \mu \sum_{m=1}^{\infty} P(N \geq m) \]

\[= \mu \mathbb{E}(N) \]

\text{Proof: } \quad N(t)+1 = m \iff N(t) = m-1

\[\iff X_1 + \ldots + X_{m-1} \leq t \]

and \[X_1 + \ldots + X_m > t\]

so that \(\{N(t)+1 = m\}\) depends only on \(X_1, \ldots, X_m\) and is therefore independent of \(X_{m+1}, X_{m+2}, \ldots\). It follows that \(N(t)+1\) is a stopping time so that

\[E(S_{N(t)+1}) = \mu E(N(t)+1)\]

\[= \mu (m(t)+1)\]
Elementary Renewal Theorem

\[
\frac{m(t)}{t} \to \frac{1}{\mu}
\]

(if \( \mu = \infty \) then \( m(t)/t \to 0 \))

**Proof.** Suppose \( \mu < \infty \). We have

\[
S_{N(t)+1} > t
\]

\[
\Rightarrow \mu(m(t)+1) > t
\]

\[
\Rightarrow \lim_{t \to \infty} \frac{m(t)}{t} \geq \frac{1}{\mu}
\]

Define truncated rv's \( X^{(M)}_m \) by

\[
X^{(M)}_m = X_m, \quad X_m \leq M
\]

\[
= M, \quad X_m > M
\]

and from the new renewal process \( N^{(M)}(t) \) using them. Clearly

\[
S_{N^{(M)}(t)+1} \leq t + M
\]
so that
\[ m_M(m_M(t) + 1) \leq t + M, \]
where \( m_M = E(X_M(t)) \) and \( m_M(t) = E(N(t)) \).

So,
\[
\lim_{t \to \infty} \frac{m_M(t)}{t} \leq \frac{1}{\mu_M}
\]

Since \( N_M(t) \geq N(t) \) we also have \( m_M(t) \geq m(t) \) and so
\[
\lim_{t \to \infty} \frac{m(t)}{t} \leq \frac{1}{\mu_M}
\]

Now let \( M \to \infty \) the MCT to get \( \mu_M \to \mu \) and
\[
\lim_{t \to \infty} \frac{m(t)}{t} \leq \frac{1}{\mu}
\]

It follows that
\[
\lim_{t \to \infty} \frac{m(t)}{t} = \lim_{t \to \infty} \frac{m(t)}{t} = \frac{1}{\mu}
\]
so that \( \frac{m(t)}{t} \to \frac{1}{\mu} \).

The case \( \mu = \infty \) is similar.